wwPDB NMR Structure Validation Summary Report

Jun 6, 2023 – 08:11 pm BST

PDB ID	:	5NKO
BMRB ID	:	34120
Title	:	Solution structure of the C-terminal domain of S. aureus Hibernating Promoting Factor (CTD-SaHPF)
Authors	:	Usachev, K.S.; Khusainov, I.S.; Ayupov, R.K.; Validov, S.Z.; Kieffer, B.; Yusupov, M.M.
Deposited on	:	2017-03-31

This is a wwPDB NMR Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/NMRValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

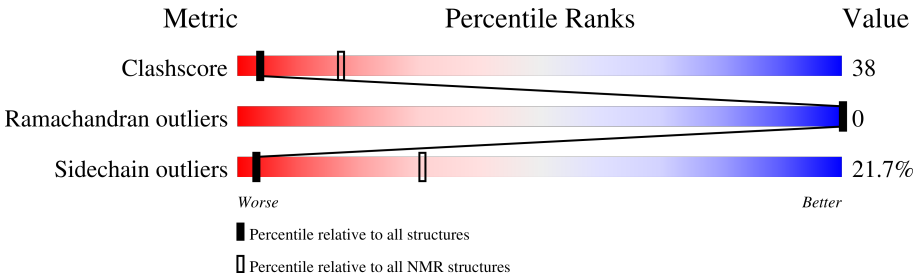
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
wwPDB-RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
wwPDB-ShiftChecker	:	v1.2
BMRB Restraints Analysis	:	v1.2
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment is 80%.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and a grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Mol	Chain	Length	Quality of chain
1	A	61	38% 54% 8%
1	B	61	33% 56% 8% .

2 Ensemble composition and analysis

This entry contains 10 models. Model 4 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	**Residue range (total)**	**Backbone RMSD (Å)**	**Medoid model**
1	A:130-A:190, B:130-B:188 (120)	1.01	4

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 2 clusters and 5 single-model clusters were found.

Cluster number	Models
1	1, 3, 4
2	8, 9
Single-model clusters	2; 5; 6; 7; 10

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 1988 atoms, of which 990 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Ribosome hibernation promotion factor.

Mol	Chain	Residues	Atoms						Trace
1	A	61	Total	C	H	N	O	S	0
			994	314	495	83	99	3	
1	B	61	Total	C	H	N	O	S	0
			994	314	495	83	99	3	

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	130	MET	-	initiating methionine	UNP Q2G055
B	130	MET	-	initiating methionine	UNP Q2G055

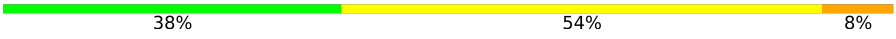
4 Residue-property plots

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

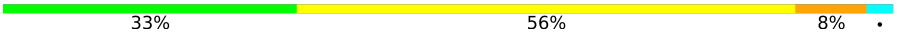
• Molecule 1: Ribosome hibernation promotion factor

Chain A:



• Molecule 1: Ribosome hibernation promotion factor

Chain B:

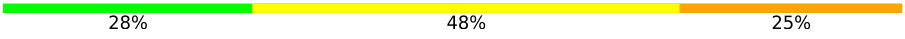


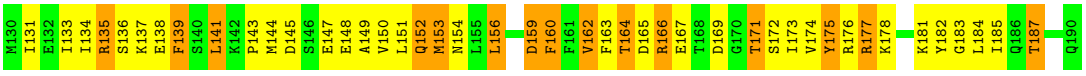
4.2 Residue scores for the representative (medoid) model from the NMR ensemble

The representative model is number 4. Colouring as in section 4.1 above.

• Molecule 1: Ribosome hibernation promotion factor

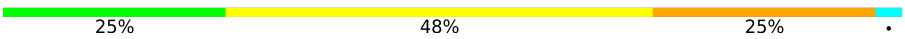
Chain A:

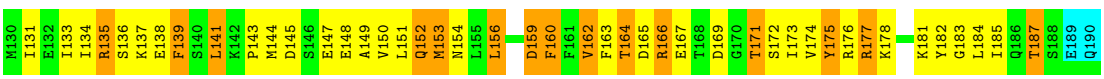




• Molecule 1: Ribosome hibernation promotion factor

Chain B:





5 Refinement protocol and experimental data overview

The models were refined using the following method: *simulated annealing.*

Of the 500 calculated structures, 10 were deposited, based on the following criterion: *structures with the lowest energy.*

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CNS	structure calculation	
CNS	refinement	

The following table shows chemical shift validation statistics as aggregates over all chemical shift files. Detailed validation can be found in section 7 of this report.

Chemical shift file(s)	working_cs.cif
Number of chemical shift lists	1
Total number of shifts	1388
Number of shifts mapped to atoms	1388
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Assignment completeness (well-defined parts)	80%

6 Model quality (i)

6.1 Standard geometry (i)

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the (average) root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#Z>5	RMSZ	#Z>5
1	A	0.58±0.04	0±0/506 (0.0± 0.0%)	0.71±0.05	0±0/675 (0.0± 0.0%)
1	B	0.59±0.04	0±0/488 (0.0± 0.0%)	0.72±0.05	0±0/654 (0.0± 0.0%)
All	All	0.59	0/9940 (0.0%)	0.72	2/13290 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	Chirality	Planarity
1	A	0.0±0.0	0.4±0.9
1	B	0.0±0.0	0.4±0.9
All	All	0	8

There are no bond-length outliers.

All unique angle outliers are listed below. They are sorted according to the Z-score of the worst occurrence in the ensemble.

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})	Models	
								Worst	Total
1	A	140	SER	CA-C-N	-5.54	105.00	117.20	2	1
1	B	140	SER	CA-C-N	-5.48	105.14	117.20	2	1

There are no chirality outliers.

5 of 8 unique planar outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Group	Models (Total)
1	A	166	ARG	Sidechain	1
1	B	166	ARG	Sidechain	1
1	A	139	PHE	Peptide	1
1	A	141	LEU	Peptide	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group	Models (Total)
1	A	142	LYS	Peptide	1

6.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in each chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	499	495	491	44±9
1	B	480	481	477	44±8
All	All	9790	9760	9680	747

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 38.

5 of 293 unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:173:ILE:HD12	1:B:173:ILE:HD12	1.10	1.24	1	5
1:A:141:LEU:HD13	1:A:162:VAL:HG22	0.97	1.37	7	1
1:B:141:LEU:HD13	1:B:162:VAL:HG22	0.92	1.37	7	1
1:A:171:THR:HG21	1:B:171:THR:HG21	0.88	1.45	4	1
1:B:151:LEU:HD13	1:B:152:GLN:N	0.84	1.87	8	4

6.3 Torsion angles (i)

6.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the backbone conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	59/61 (97%)	56±1 (94±3%)	3±1 (6±3%)	0±0 (0±0%)	100	100
1	B	58/61 (95%)	55±1 (95±2%)	3±1 (5±2%)	0±0 (0±0%)	100	100
All	All	1170/1220 (96%)	1106 (95%)	64 (5%)	0 (0%)	100	100

There are no Ramachandran outliers.

6.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all NMR entries. The Analysed column shows the number of residues for which the sidechain conformation was analysed and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	56/56 (100%)	44±3 (78±6%)	12±3 (22±6%)	3 31
1	B	54/56 (96%)	42±3 (78±6%)	12±3 (22±6%)	3 30
All	All	1100/1120 (98%)	861 (78%)	239 (22%)	3 30

5 of 64 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	166	ARG	10
1	A	175	TYR	10
1	B	166	ARG	10
1	B	175	TYR	10
1	A	156	LEU	7

6.3.3 RNA

There are no RNA molecules in this entry.

6.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.5 Carbohydrates

There are no monosaccharides in this entry.

6.6 Ligand geometry

There are no ligands in this entry.

6.7 Other polymers (i)

There are no such molecules in this entry.

6.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

7 Chemical shift validation

The completeness of assignment taking into account all chemical shift lists is 80% for the well-defined parts and 81% for the entire structure.

7.1 Chemical shift list 1

File name: working_cs.cif

Chemical shift list name: *assigned_chem_shift_list*

7.1.1 Bookkeeping

The following table shows the results of parsing the chemical shift list and reports the number of nuclei with statistically unusual chemical shifts.

Total number of shifts	1388
Number of shifts mapped to atoms	1388
Number of unparsed shifts	0
Number of shifts with mapping errors	0
Number of shifts with mapping warnings	0
Number of shift outliers (ShiftChecker)	0

7.1.2 Chemical shift referencing

The following table shows the suggested chemical shift referencing corrections.

Nucleus	# values	Correction ± precision, *ppm*	Suggested action
$^{13}C_{\alpha}$	120	-0.13 ± 0.10	None needed (< 0.5 ppm)
$^{13}C_{\beta}$	112	-0.07 ± 0.15	None needed (< 0.5 ppm)
$^{13}C'$	116	0.01 ± 0.18	None needed (< 0.5 ppm)
^{15}N	112	0.10 ± 0.56	None needed (< 0.5 ppm)

7.1.3 Completeness of resonance assignments

The following table shows the completeness of the chemical shift assignments for the well-defined regions of the structure. The overall completeness is 80%, i.e. 1361 atoms were assigned a chemical shift out of a possible 1691. 0 out of 16 assigned methyl groups (LEU and VAL) were assigned stereospecifically.

	Total	^{1}H	^{13}C	^{15}N
Backbone	576/604 (95%)	234/246 (95%)	232/240 (97%)	110/118 (93%)
Sidechain	711/955 (74%)	462/616 (75%)	242/300 (81%)	7/39 (18%)

Continued on next page...

Continued from previous page...

	Total	^{1}H	^{13}C	^{15}N
Aromatic	74/132 (56%)	40/64 (62%)	34/64 (53%)	0/4 (0%)
Overall	1361/1691 (80%)	736/926 (79%)	508/604 (84%)	117/161 (73%)

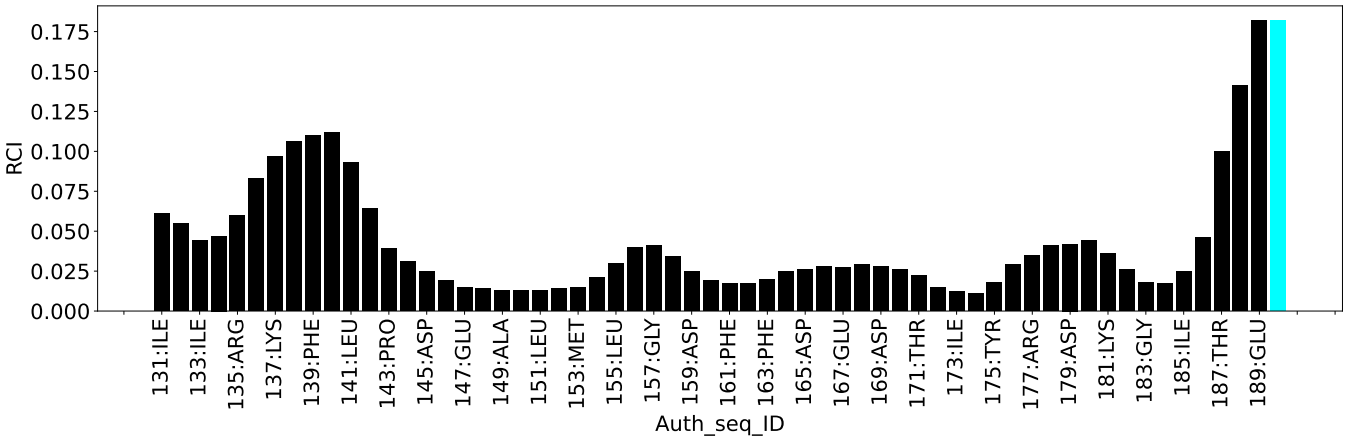
7.1.4 Statistically unusual chemical shifts (i)

There are no statistically unusual chemical shifts.

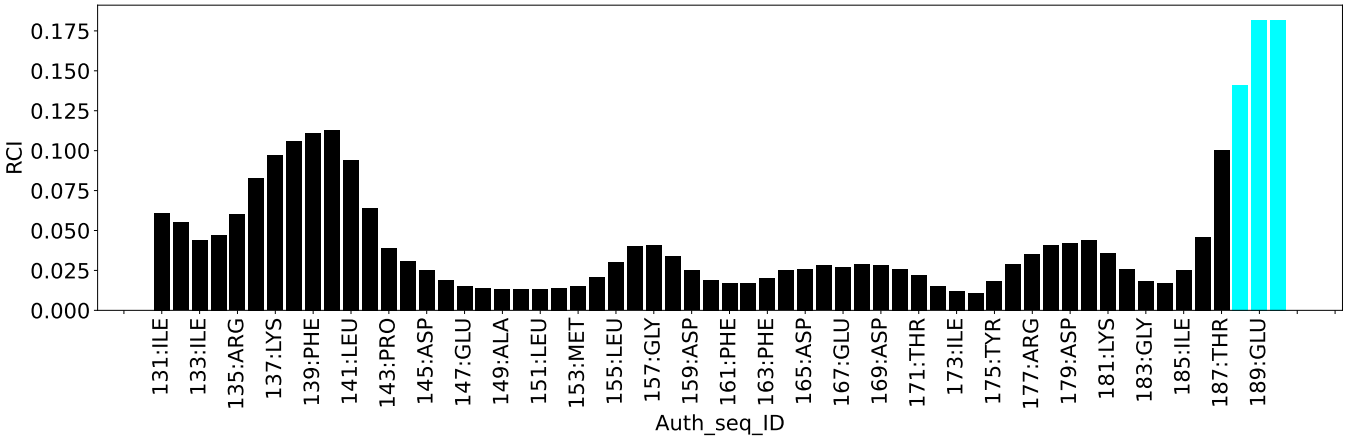
7.1.5 Random Coil Index (RCI) plots (i)

The image below reports *random coil index* values for the protein chains in the structure. The height of each bar gives a probability of a given residue to be disordered, as predicted from the available chemical shifts and the amino acid sequence. A value above 0.2 is an indication of significant predicted disorder. The colour of the bar shows whether the residue is in the well-defined core (black) or in the ill-defined residue ranges (cyan), as described in section 2 on ensemble composition.If well-defined core and ill-defined regions are not identified then it is shown as gray bars.

Random coil index (RCI) for chain A:



Random coil index (RCI) for chain B:


RCI
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
131:ILE
133:ILE
135:ARG
137:LYS
139:PHE
141:LEU
143:PRO
145:ASP
147:GLU
149:ALA
151:LEU
153:MET
155:LEU
157:GLY
159:ASP
161:PHE
163:PHE
165:ASP
167:GLU
169:ASP
171:THR
173:ILE
175:TYR
177:ARG
179:ASP
181:LYS
183:GLY
185:ILE
187:THR
189:GLU
Auth_seq_ID

8 NMR restraints analysis

8.1 Conformationally restricting restraints

The following table provides the summary of experimentally observed NMR restraints in different categories. Restraints are classified into different categories based on the sequence separation of the atoms involved.

Description	Value
Total distance restraints	2296
Intra-residue (|i-j|=0)	949
Sequential (|i-j|=1)	489
Medium range (|i-j|>1 and |i-j|<5)	283
Long range (|i-j|≥5)	414
Inter-chain	53
Hydrogen bond restraints	108
Disulfide bond restraints	0
Total dihedral-angle restraints	228
Number of unmapped restraints	0
Number of restraints per residue	20.7
Number of long range restraints per residue[1]	3.9

[1]Long range hydrogen bonds and disulfide bonds are counted as long range restraints while calculating the number of long range restraints per residue

8.2 Residual restraint violations

This section provides the overview of the restraint violations analysis. The violations are binned as small, medium and large violations based on its absolute value. Average number of violations per model is calculated by dividing the total number of violations in each bin by the size of the ensemble.

8.2.1 Average number of distance violations per model

Distance violations less than 0.1 Å are not included in the calculation.

Bins (Å)	Average number of violations per model	Max (Å)
0.1-0.2 (Small)	165.0	0.2
0.2-0.5 (Medium)	287.4	0.5
>0.5 (Large)	295.9	20.12

8.2.2 Average number of dihedral-angle violations per model (i)

Dihedral-angle violations less than 1° are not included in the calculation.

Bins (°)	Average number of violations per model	Max (°)
1.0-10.0 (Small)	45.0	9.4
10.0-20.0 (Medium)	0.6	13.8
>20.0 (Large)	None	None

9 Distance violation analysis (i)

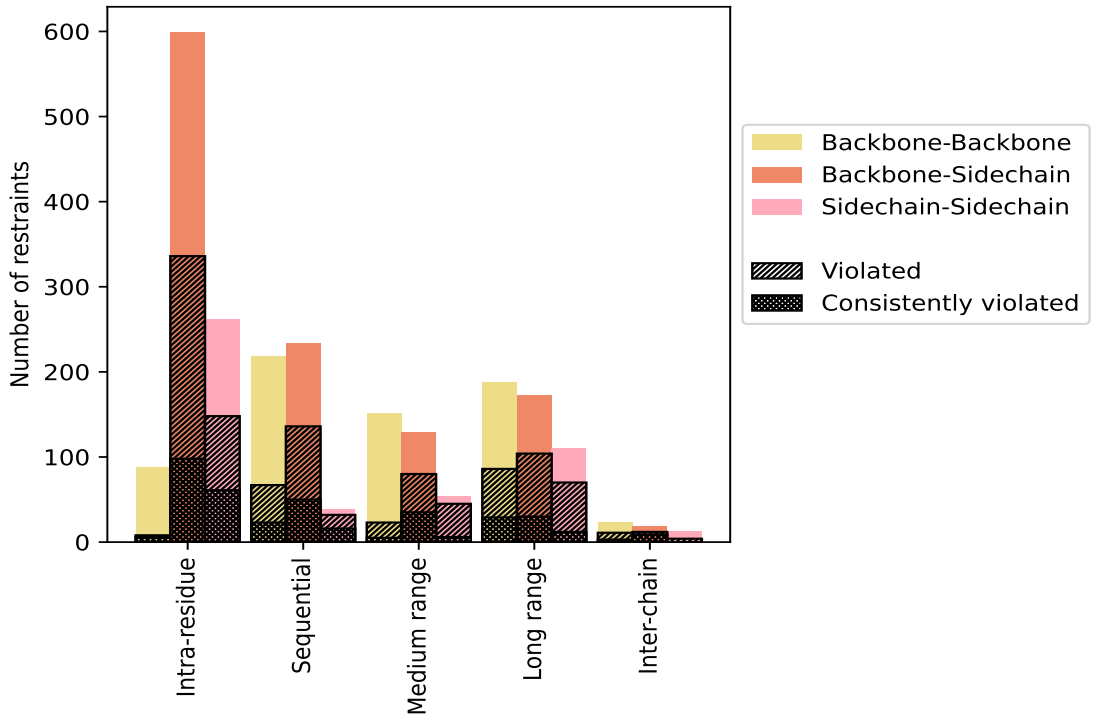
9.1 Summary of distance violations (i)

The following table shows the summary of distance violations in different restraint categories based on the sequence separation of the atoms involved. Each category is further sub-divided into three sub-categories based on the atoms involved. Violations less than 0.1 Å are not included in the statistics.

Restraints type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
Intra-residue (|i-j|=0)	949	41.3	492	51.8	21.4	165	17.4	7.2
Backbone-Backbone	88	3.8	8	9.1	0.3	6	6.8	0.3
Backbone-Sidechain	599	26.1	336	56.1	14.6	98	16.4	4.3
Sidechain-Sidechain	262	11.4	148	56.5	6.4	61	23.3	2.7
Sequential (|i-j|=1)	489	21.3	235	48.1	10.2	89	18.2	3.9
Backbone-Backbone	218	9.5	67	30.7	2.9	23	10.6	1.0
Backbone-Sidechain	233	10.1	136	58.4	5.9	50	21.5	2.2
Sidechain-Sidechain	38	1.7	32	84.2	1.4	16	42.1	0.7
Medium range (|i-j|>1 & |i-j|<5)	283	12.3	144	50.9	6.3	46	16.3	2.0
Backbone-Backbone	100	4.4	19	19.0	0.8	5	5.0	0.2
Backbone-Sidechain	129	5.6	80	62.0	3.5	35	27.1	1.5
Sidechain-Sidechain	54	2.4	45	83.3	2.0	6	11.1	0.3
Long range (|i-j|≥5)	414	18.0	255	61.6	11.1	71	17.1	3.1
Backbone-Backbone	132	5.7	81	61.4	3.5	29	22.0	1.3
Backbone-Sidechain	172	7.5	104	60.5	4.5	30	17.4	1.3
Sidechain-Sidechain	110	4.8	70	63.6	3.0	12	10.9	0.5
Inter-chain	53	2.3	26	49.1	1.1	11	20.8	0.5
Backbone-Backbone	22	1.0	10	45.5	0.4	2	9.1	0.1
Backbone-Sidechain	19	0.8	12	63.2	0.5	9	47.4	0.4
Sidechain-Sidechain	12	0.5	4	33.3	0.2	0	0.0	0.0
Hydrogen bond	108	4.7	10	9.3	0.4	1	0.9	0.0
Disulfide bond	0	0.0	0	0.0	0.0	0	0.0	0.0
Total	2296	100.0	1162	50.6	50.6	383	16.7	16.7
Backbone-Backbone	668	29.1	195	29.2	8.5	66	9.9	2.9
Backbone-Sidechain	1152	50.2	668	58.0	29.1	222	19.3	9.7
Sidechain-Sidechain	476	20.7	299	62.8	13.0	95	20.0	4.1

[1] percentage calculated with respect to the total number of distance restraints, [2] percentage calculated with respect to the number of restraints in a particular restraint category, [3] violated in at least one model, [4] violated in all the models

9.1.1 Bar chart : Distribution of distance restraints and violations



Violated and consistently violated restraints are shown using different hatch patterns in their respective categories. The hydrogen bonds and disulfied bonds are counted in their appropriate category on the x-axis

9.2 Distance violation statistics for each model

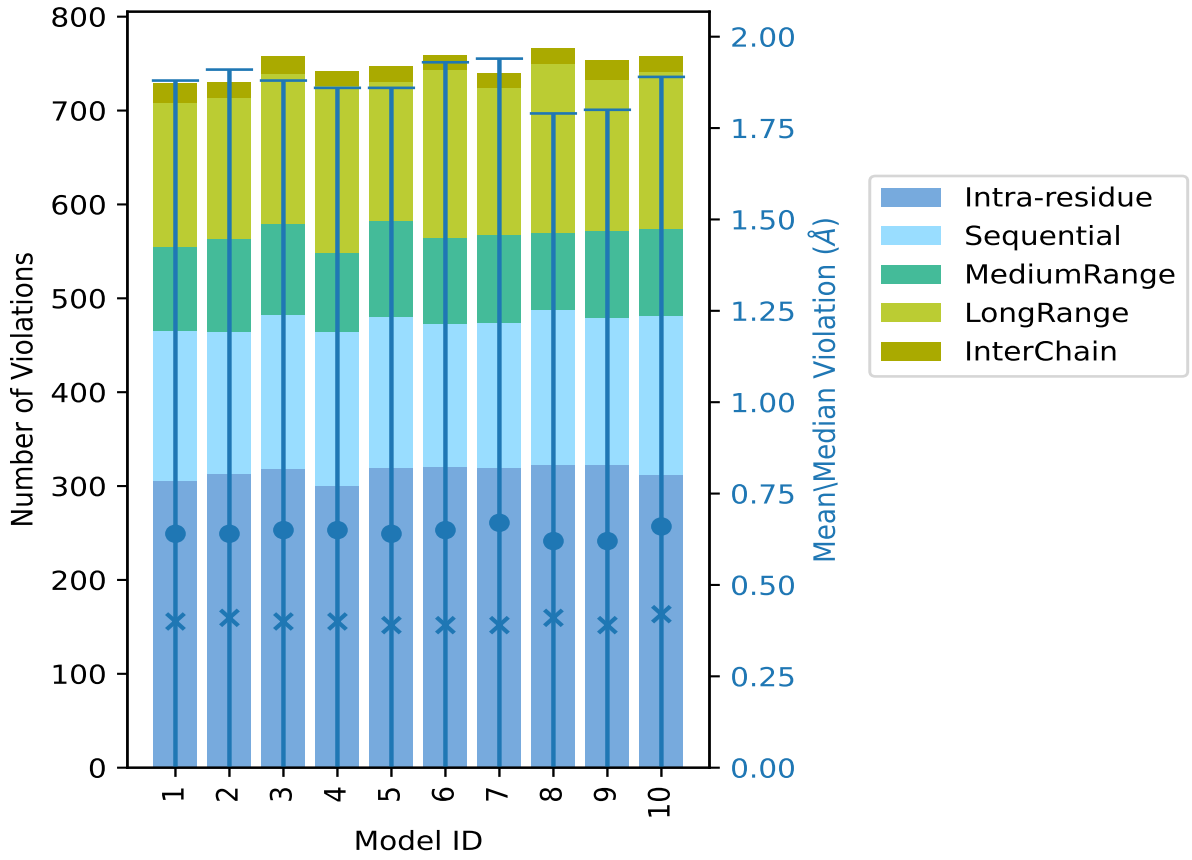
The following table provides the distance violation statistics for each model in the ensemble. Violations less than 0.1 Å are not included in the statistics.

Model ID	Number of violations						Mean (Å)	Max (Å)	SD[6] (Å)	Median (Å)
	IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total				
1	306	159	90	153	21	729	0.64	19.36	1.24	0.4
2	313	151	99	150	17	730	0.64	20.12	1.27	0.41
3	319	163	98	159	19	758	0.65	19.28	1.23	0.4
4	301	163	85	178	15	742	0.65	18.46	1.21	0.4
5	320	160	102	148	17	747	0.64	19.67	1.22	0.39
6	320	153	91	180	15	759	0.65	18.62	1.28	0.39
7	319	155	94	156	15	739	0.67	19.24	1.27	0.39
8	322	166	82	180	17	767	0.62	18.73	1.17	0.41
9	323	156	93	161	21	754	0.62	18.95	1.18	0.39
10	312	170	92	168	16	758	0.66	18.65	1.23	0.42

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints,

[5]Inter-chain restraints, [6]Standard deviation

9.2.1 Bar graph : Distance Violation statistics for each model



The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

9.3 Distance violation statistics for the ensemble

Violation analysis may find that some restraints are violated in few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of the ensemble. In total, 1036(IR:457, SQ:254, MR:139, LR:159, IC:27) restraints are not violated in the ensemble.

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
55	9	15	29	1	109	1	10.0
41	21	19	17	5	103	2	20.0
44	23	7	10	5	89	3	30.0
33	27	9	19	1	89	4	40.0

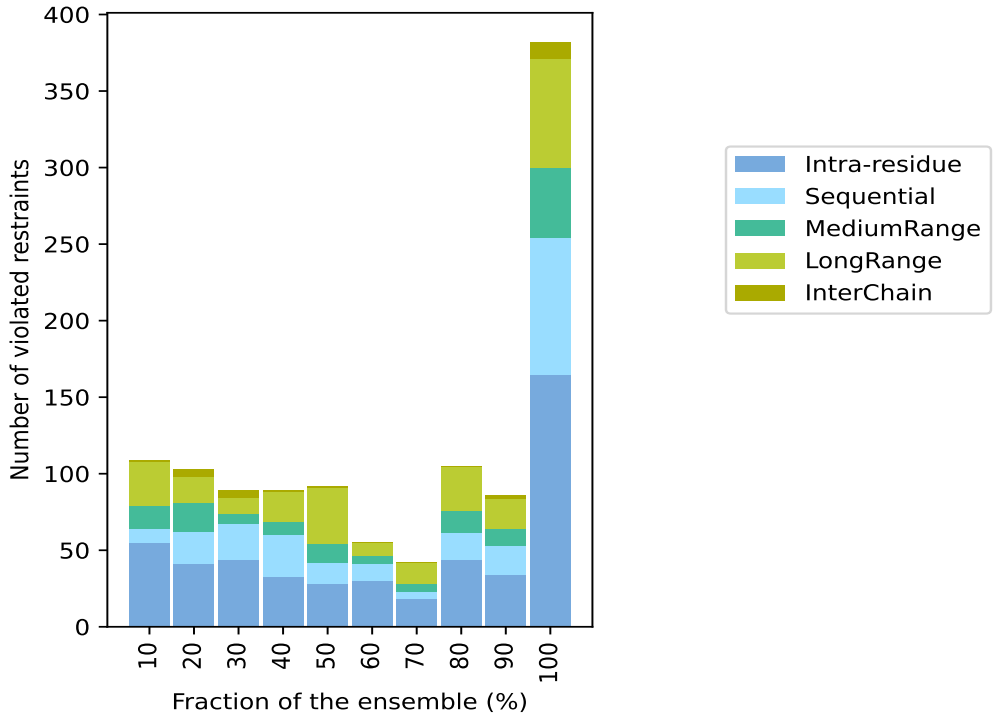
Continued on next page...

Continued from previous page...

Number of violated restraints						Fraction of the ensemble	
IR[1]	SQ[2]	MR[3]	LR[4]	IC[5]	Total	Count[6]	%
28	14	12	37	1	92	5	50.0
30	11	5	9	0	55	6	60.0
18	5	5	14	0	42	7	70.0
44	17	15	29	0	105	8	80.0
34	19	11	20	2	86	9	90.0
165	89	46	71	11	382	10	100.0

[1]Intra-residue restraints, [2]Sequential restraints, [3]Medium range restraints, [4]Long range restraints, [5]Inter-chain restraints, [6] Number of models with violations

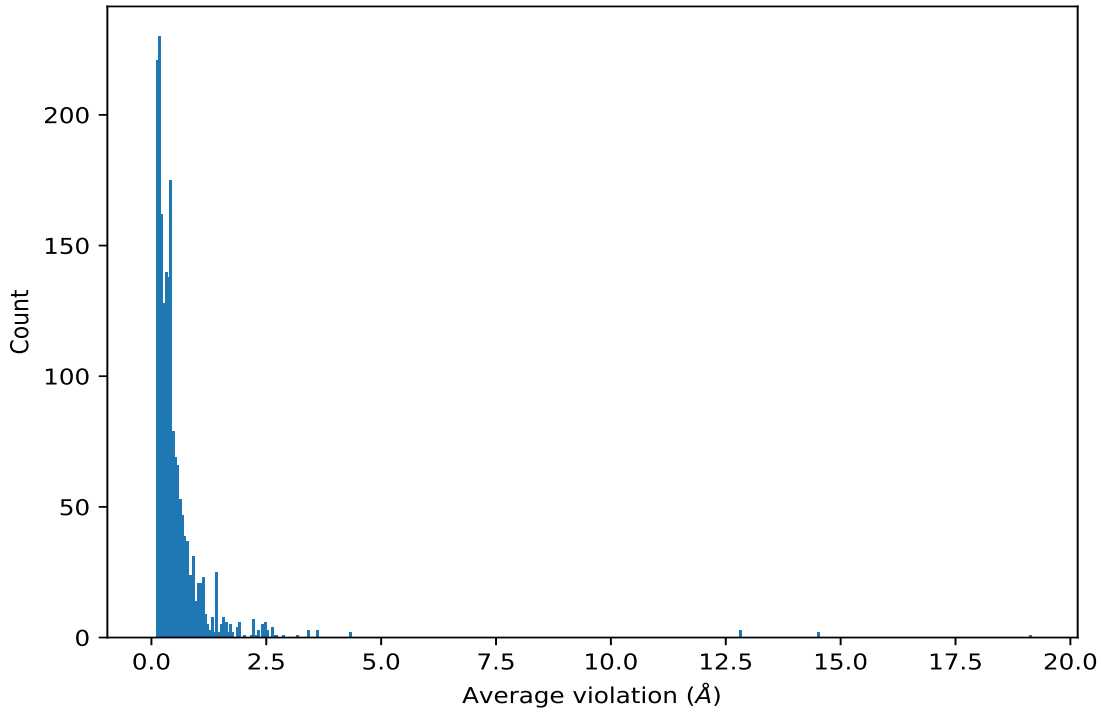
9.3.1 Bar graph : Distance violation statistics for the ensemble



9.4 Most violated distance restraints in the ensemble

9.4.1 Histogram : Distribution of mean distance violations

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models in the ensemble



9.4.2 Table: Most violated distance restraints (i)

The following table provides the mean and the standard deviation of the violations for the 10 worst performing restraints, sorted by number of violated models and the mean violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

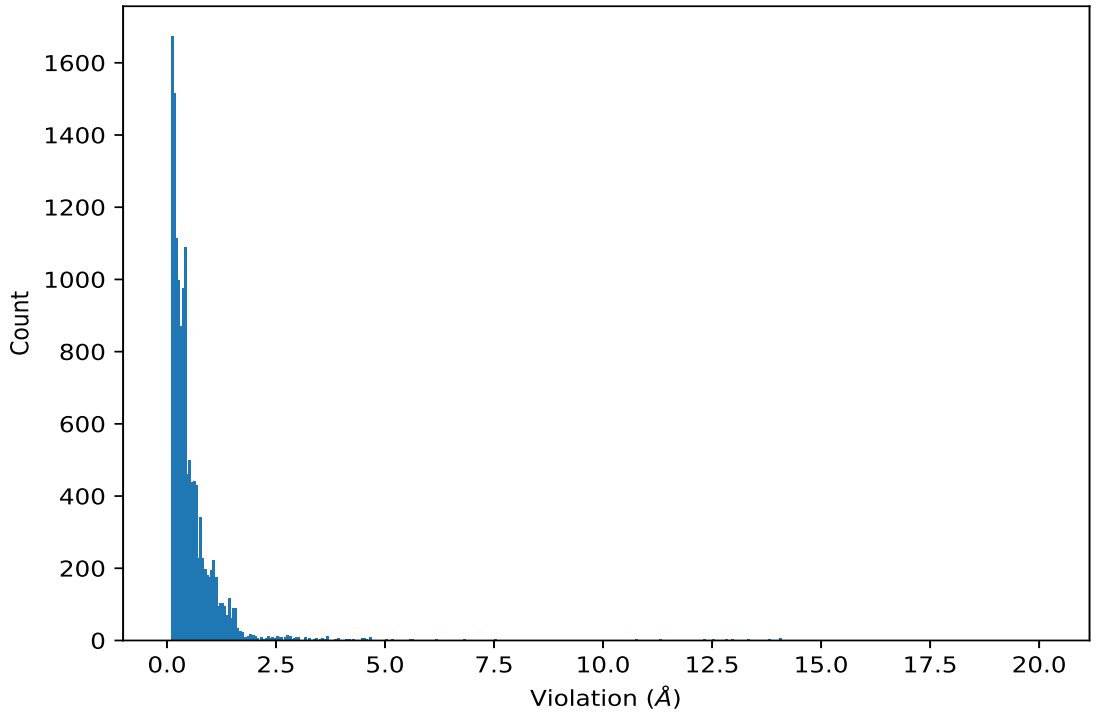
Key	Atom-1	Atom-2	Models[1]	Mean (Å)	SD[1] (Å)	Median (Å)
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	10	19.11	0.5	19.1
(2,957)	1:B:134:ILE:H	1:B:183:GLY:H	10	14.53	0.73	14.31
(2,958)	1:A:134:ILE:H	1:A:183:GLY:H	10	14.53	0.73	14.32
(2,874)	1:B:158:HIS:H	1:A:156:LEU:HD21	10	12.8	1.06	12.9
(2,874)	1:B:158:HIS:H	1:A:156:LEU:HD22	10	12.8	1.06	12.9
(2,874)	1:B:158:HIS:H	1:A:156:LEU:HD23	10	12.8	1.06	12.9
(2,632)	1:A:162:VAL:H	1:A:141:LEU:HD11	10	3.43	1.97	3.06
(2,632)	1:A:162:VAL:H	1:A:141:LEU:HD12	10	3.43	1.97	3.06
(2,632)	1:A:162:VAL:H	1:A:141:LEU:HD13	10	3.43	1.97	3.06
(2,1961)	1:B:176:ARG:HG3	1:B:178:LYS:HB2	10	3.15	1.22	3.26
(2,479)	1:A:142:LYS:H	1:A:143:PRO:HB2	10	2.87	0.06	2.88
(2,530)	1:A:153:MET:H	1:A:151:LEU:HG	10	2.7	0.28	2.65
(2,529)	1:B:153:MET:H	1:B:151:LEU:HG	10	2.69	0.28	2.65
(2,1451)	1:B:156:LEU:HA	1:B:155:LEU:HD11	10	2.53	0.69	2.78

[1]Number of violated models, [2]Standard deviation

9.5 All violated distance restraints (i)

9.5.1 Histogram : Distribution of distance violations (i)

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



9.5.2 Table : All distance violations (i)

The following table provides the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint. Rows with same key represent combinatorial or ambiguous restraints and are counted as a single restraint.

Key	Atom-1	Atom-2	Model ID	Violation (Å)
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	2	20.12
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	5	19.67
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	1	19.36
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	3	19.28
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	7	19.24
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	9	18.95
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	8	18.73
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	10	18.65
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	6	18.62
(1,22)	1:A:150:VAL:N	1:B:146:SER:O	4	18.46

10 Dihedral-angle violation analysis

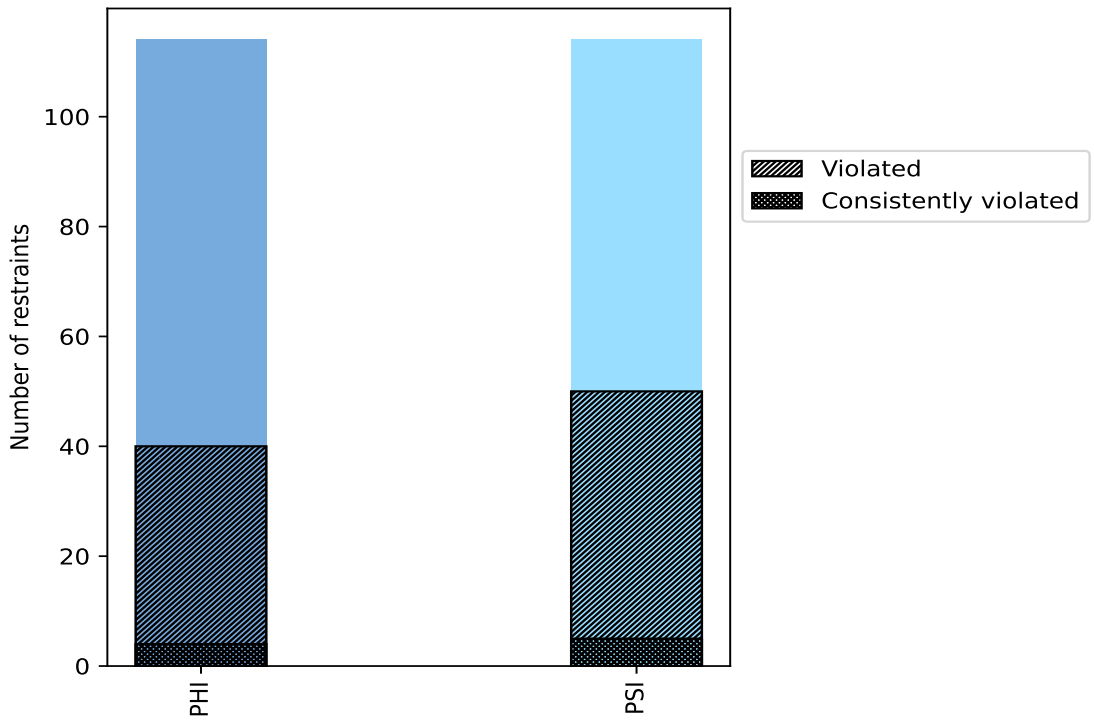
10.1 Summary of dihedral-angle violations

The following table provides the summary of dihedral-angle violations in different dihedral-angle types. Violations less than 1° are not included in the calculation.

Angle type	Count	%[1]	Violated[3]			Consistently Violated[4]		
			Count	%[2]	%[1]	Count	%[2]	%[1]
PHI	114	50.0	40	35.1	17.5	4	3.5	1.8
PSI	114	50.0	50	43.9	21.9	5	4.4	2.2
Total	228	100.0	90	39.5	39.5	9	3.9	3.9

[1] percentage calculated with respect to total number of dihedral-angle restraints, [2] percentage calculated with respect to number of restraints in a particular dihedral-angle type, [3] violated in at least one model, [4] violated in all the models

10.1.1 Bar chart : Distribution of dihedral-angles and violations



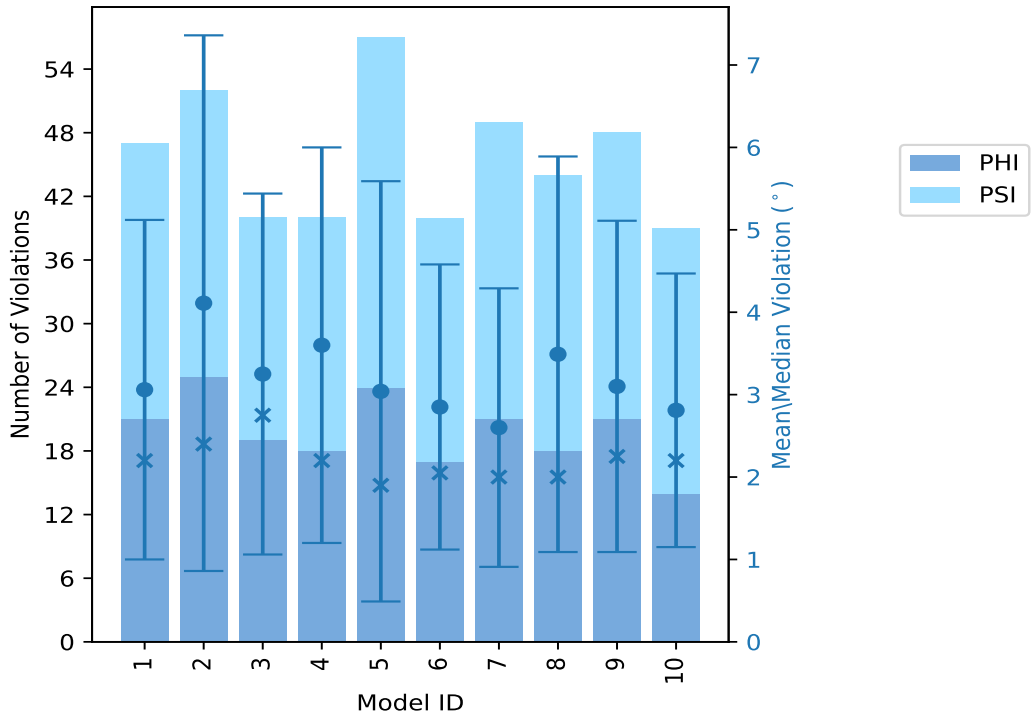
Violated and consistently violated restraints are shown using different hatch patterns in their respective categories

10.2 Dihedral-angle violation statistics for each model (i)

The following table provides the dihedral-angle violation statistics for each model in the ensemble. Violations less than 1° are not included in the statistics.

Model ID	Number of violations			Mean (°)	Max (°)	SD (°)	Median (°)
	PHI	PSI	Total				
1	21	26	47	3.06	9.1	2.06	2.2
2	25	27	52	4.11	13.8	3.25	2.4
3	19	21	40	3.25	9.4	2.19	2.75
4	18	22	40	3.6	8.6	2.4	2.2
5	24	33	57	3.04	12.5	2.55	1.9
6	17	23	40	2.85	6.6	1.73	2.05
7	21	28	49	2.6	8.5	1.69	2.0
8	18	26	44	3.49	8.1	2.4	2.0
9	21	27	48	3.1	8.0	2.01	2.25
10	14	25	39	2.81	7.4	1.66	2.2

10.2.1 Bar graph : Dihedral violation statistics for each model (i)



The mean(dot),median(x) and the standard deviation are shown in blue with respect to the y axis on the right

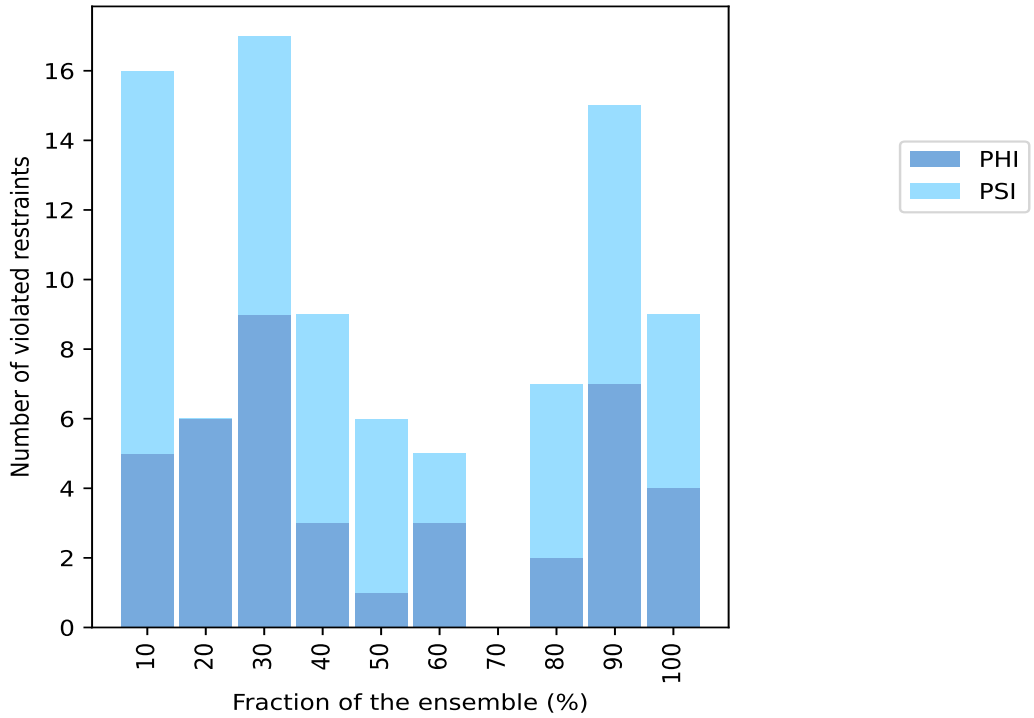
10.3 Dihedral-angle violation statistics for the ensemble (i)

Violation analysis may find that some restraints are violated in very few models and some are violated in most of models. The following table provides this information as number of violated restraints for a given fraction of ensemble.

Number of violated restraints			Fraction of the ensemble	
PHI	PSI	Total	Count[1]	%
5	11	16	1	10.0
6	0	6	2	20.0
9	8	17	3	30.0
3	6	9	4	40.0
1	5	6	5	50.0
3	2	5	6	60.0
0	0	0	7	70.0
2	5	7	8	80.0
7	8	15	9	90.0
4	5	9	10	100.0

[1] Number of models with violations

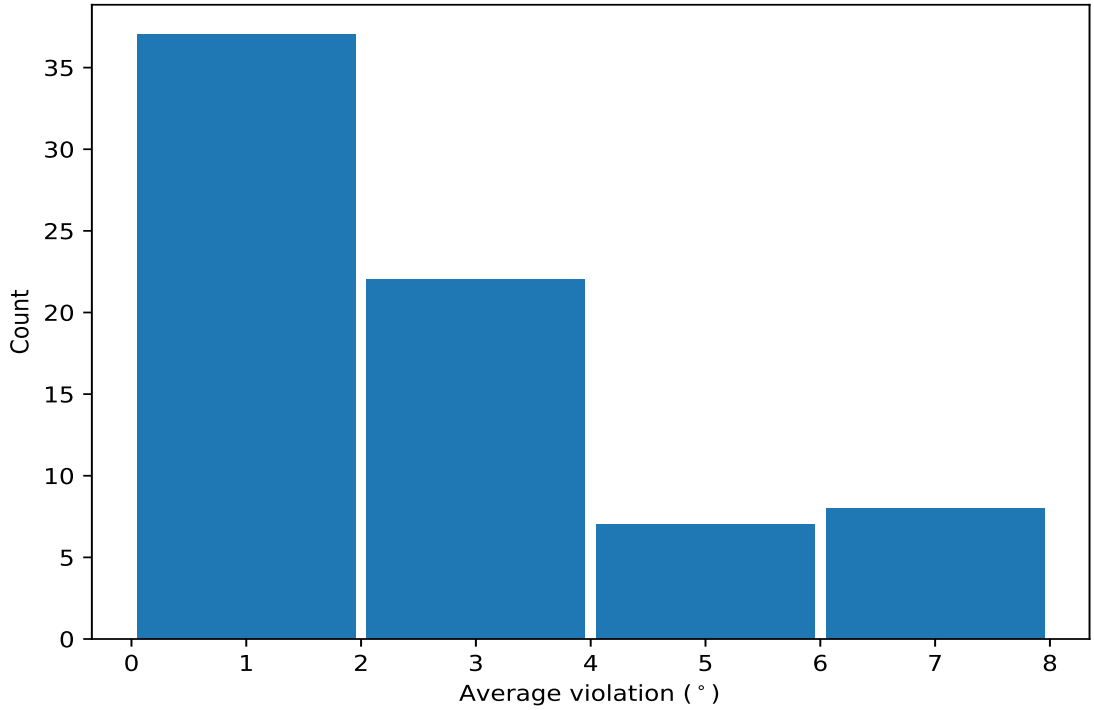
10.3.1 Bar graph : Dihedral-angle Violation statistics for the ensemble (i)



10.4 Most violated dihedral-angle restraints in the ensemble

10.4.1 Histogram : Distribution of mean dihedral-angle violations

The following histogram shows the distribution of the average value of the violation. The average is calculated for each restraint that is violated in more than one model over all the violated models in the ensemble



10.4.2 Table: Most violated dihedral-angle restraints

The following table provides the mean and the standard deviation of the violations for the 10 worst performing restraints, sorted by number of violated models and the mean violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

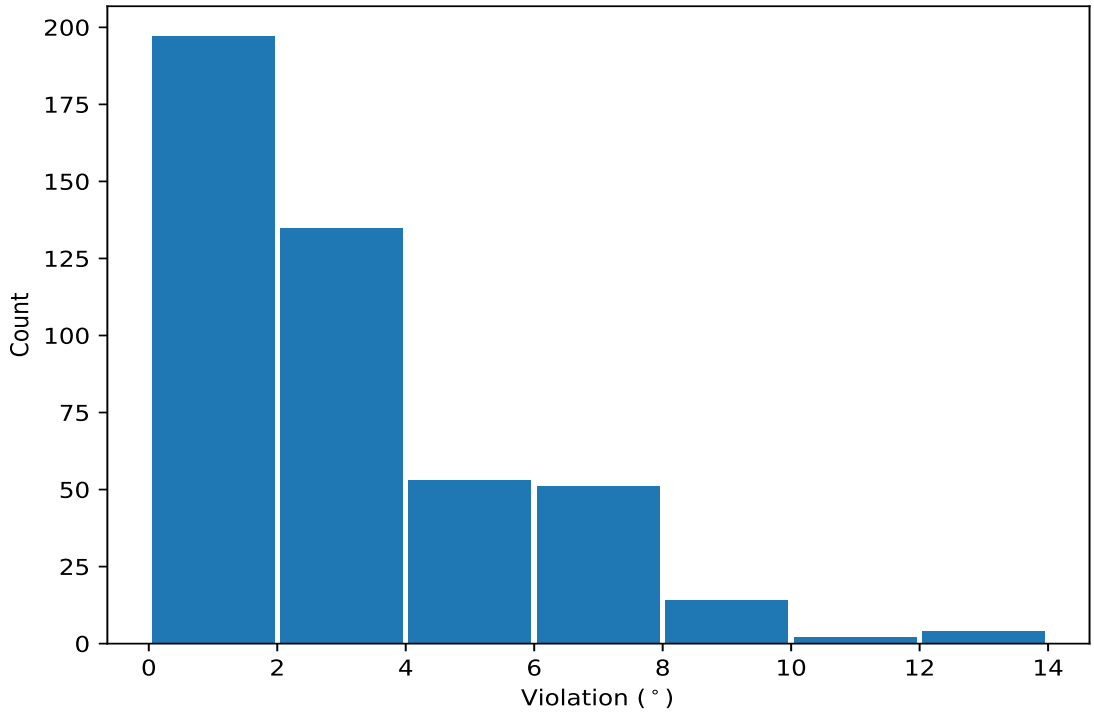
Key	Atom-1	Atom-2	Atom-3	Atom-4	Models[1]	Mean	SD[2]	Median
(1,132)	1:B:140:SER:N	1:B:140:SER:CA	1:B:140:SER:C	1:B:141:LEU:N	10	6.27	2.74	5.95
(1,18)	1:A:140:SER:N	1:A:140:SER:CA	1:A:140:SER:C	1:A:141:LEU:N	10	6.26	2.75	6.0
(1,128)	1:B:138:GLU:N	1:B:138:GLU:CA	1:B:138:GLU:C	1:B:139:PHE:N	10	6.24	1.78	6.6
(1,14)	1:A:138:GLU:N	1:A:138:GLU:CA	1:A:138:GLU:C	1:A:139:PHE:N	10	5.96	1.74	6.45
(1,19)	1:A:140:SER:C	1:A:141:LEU:N	1:A:141:LEU:CA	1:A:141:LEU:C	10	5.31	1.94	5.35
(1,133)	1:B:140:SER:C	1:B:141:LEU:N	1:B:141:LEU:CA	1:B:141:LEU:C	10	5.3	1.98	5.35
(1,115)	1:B:131:ILE:C	1:B:132:GLU:N	1:B:132:GLU:CA	1:B:132:GLU:C	10	3.59	0.89	3.65
(1,1)	1:A:131:ILE:C	1:A:132:GLU:N	1:A:132:GLU:CA	1:A:132:GLU:C	10	3.56	0.86	3.7
(1,26)	1:A:144:MET:N	1:A:144:MET:CA	1:A:144:MET:C	1:A:145:ASP:N	10	2.27	0.77	2.25
(1,131)	1:B:139:PHE:C	1:B:140:SER:N	1:B:140:SER:CA	1:B:140:SER:C	9	6.67	2.65	5.9

[1] Number of violated models, [2]Standard deviation, All angle values are in degree (°)

10.5 All violated dihedral-angle restraints

10.5.1 Histogram : Distribution of violations

The following histogram shows the distribution of the absolute value of the violation for all violated restraints in the ensemble.



10.5.2 Table: All violated dihedral-angle restraints

The following table provides the list of violationsfor the 10 worst performing restraints, sorted by the violation value. The Key (restraint list ID, restraint ID) is the unique identifier for a given restraint.

Key	Atom-1	Atom-2	Atom-3	Atom-4	Model ID	Violation (°)
(1,17)	1:A:139:PHE:C	1:A:140:SER:N	1:A:140:SER:CA	1:A:140:SER:C	2	13.8
(1,131)	1:B:139:PHE:C	1:B:140:SER:N	1:B:140:SER:CA	1:B:140:SER:C	2	13.7
(1,18)	1:A:140:SER:N	1:A:140:SER:CA	1:A:140:SER:C	1:A:141:LEU:N	5	12.5
(1,132)	1:B:140:SER:N	1:B:140:SER:CA	1:B:140:SER:C	1:B:141:LEU:N	5	12.5
(1,21)	1:A:141:LEU:C	1:A:142:LYS:N	1:A:142:LYS:CA	1:A:142:LYS:C	2	10.5
(1,135)	1:B:141:LEU:C	1:B:142:LYS:N	1:B:142:LYS:CA	1:B:142:LYS:C	2	10.4
(1,130)	1:B:139:PHE:N	1:B:139:PHE:CA	1:B:139:PHE:C	1:B:140:SER:N	3	9.4
(1,16)	1:A:139:PHE:N	1:A:139:PHE:CA	1:A:139:PHE:C	1:A:140:SER:N	3	9.3
(1,19)	1:A:140:SER:C	1:A:141:LEU:N	1:A:141:LEU:CA	1:A:141:LEU:C	5	9.2
(1,133)	1:B:140:SER:C	1:B:141:LEU:N	1:B:141:LEU:CA	1:B:141:LEU:C	5	9.2